

Detection of a watermark in a digital signal

The present invention relates to the detection of watermark signals embedded in digital data, the data typically representing multimedia content. A typical format for such data is MPEG2, although the invention may be used with other formats also.

5

In order to embed certain information, such as copyright, copy control, source or authentication data into a digital signal, a technique known as watermarking is often used. This involves processing the digital data so that a recognizable pattern is 'overlaid' onto the data to be watermarked. Different types of watermark have different uses. A simple robust
10 watermark, which is intended to survive a wide range of processing steps in the analogue and digital domains, may simply indicate that the watermarked data is subject to copyright, and may provide further details, such as owner and date. A fragile watermark is often added in such a way that it is corrupted or broken if the data is processed in any way. In this way, the absence of a fragile watermark in a data file, or stream, in which one was expected, can
15 indicate that the data has been processed or otherwise tampered with. This can be useful in medical or forensic science applications where authenticity is crucial.

The various types of watermark pattern themselves consist of a pseudo-noise signal which is overlaid onto, or woven into, the data itself. The watermark signal should ideally not degrade the source data in a perceptible manner, but should be detectable by a
20 suitable decoder.

A particular problem arises when the watermarked data is compressed to a very low bit-rate, suitable for transmission over the Internet, or other data transfer system. DIVX is one system which produces very low bit-rates, and is widely used to reduce the amount of bandwidth required to transmit video images over the Internet.

25 Currently used watermarking systems such as JAWS (Ton Kalker, Geert Depovere, Jaap Haitzma, Maurice Maes, "A Video Watermarking System for Broadcast Monitoring", Proceedings of SPIE Electronic Imaging '99, Security and watermarking of Multimedia Contents, San Jose (CA), USA, January 1999) use detectors which search for embedded watermarks by collecting large amounts of video data, which is then folded and

accumulated before the accumulated data is correlated with the expected watermark pattern. With video data that has been compressed to a very low bit-rate, e.g. using DIVX, a frequently encountered result is correlation peaks which occur below the detection threshold. This means that detection of the embedded watermark(s) may fail, which can cause
5 inconvenience for users of the system who may be authorized to view the watermarked video, but are prevented from doing so in the absence of a proper detection of the watermark(s).

A further problem occurs when the watermarked video has been scaled or re-sized. In order to detect an embedded watermark, the original scale of the video signal is
10 required, so that the accumulation buffer, which captures incoming video data, can be correspondingly scaled to the original video dimensions. The original scale must be determined from the scaled video data itself. Compared to the watermark detection process, where the video data is correlated against known watermark data, prior art scale-detection processes operate by correlating two noisy accumulation buffers with each other to yield the
15 scale factor.

In the JAWS system, watermark detection and the watermark detection process and the scale retrieval process make use of a repetitive watermark pattern being embedded in the source data. During the watermark embedding process, a 128x128 watermark pattern is 'tiled' over the full extent of a frame of data.

20 In order to retrieve the horizontal scale information from a scaled version of the data, the process begins by arbitrarily selecting two horizontally adjacent tiles A and B from a number of accumulated frames. The two tiles are then correlated with each other according to the following steps:

- Calculate 128x128 Hanning window over A and B; $Han(A)$, $Han(B)$
25 ○ A Hanning window is a kind of filter which acts to 'fade out' the edges of the tile to which it is applied. In this way, the data in the centre of the tile is preserved, but closer to the edges, the data fades to zero. This alleviates the effect of edges introducing strong artificial frequency components in the ensuing FFT calculation.
- Calculate 128x128 Fast Fourier Transform (FFT) over A and B
- 30 • Calculate complex conjugate of $Han(B)$; $Con(Han(B))$
- Calculate pointwise multiplication of $Han(A)$ and $Con(Han(B))$
- Normalise multiplication result. This is done according to the following formula for each complex value (z) in the result, so that z is replaced by

$$\frac{z}{\sqrt{re(z)^2 + im(z)^2}}$$

- Calculate Inverse FFT of previous step

The position of the highest value in the first row of the IFFT result is then used to calculate the horizontal scale factor. If the first value is the highest, then the horizontal scaling factor is 1 i.e. no scaling has occurred.

The vertical scaling factor is calculated in a similar way, but two vertically adjacent tiles and the first column of the IFFT result are used instead.

The correlation peaks for this scale retrieval process are even lower than for the watermark detection process due to the inherently more noisy buffer samples used.

- (Watermark detection involves a correlation between a known pattern and a noisy accumulation buffer: scale detection is a correlation between two noisy accumulation buffers). To further complicate matters, frame folding may not be used in the scale detection process. This is because frame folding can only be used if the scale is known. If the scale is not known, patterns are accumulated that are not synchronised and the resulting accumulation buffer is useless. As a result, only accumulation can be used. This means that more frames must be collected before correlation can be performed, which, of course, takes more time.

- Folding works by 'magnifying' the watermark data, as it always has the same sign. The underlying video signal is effectively 'random' and so averaged out. Folding for long enough results in the original watermark pattern. However, if the patterns (tile of 128x128) are not exactly aligned the process does not work.

Prior art techniques attempt to alleviate these problems by accumulating more frames per detection in the hope that the video data averages out and the watermark signal amplifies, so that the signal (watermark) to noise (video) ratio increases.

- In a typical scale-detection, up to 300 frames are currently used. However, in the case of DIVX compressed video, a lot of artificial noise and undesired similarity, caused by block patterns, is introduced. During the accumulation process, more noise than watermark energy is generally accumulated. Also, the undesired patterns are amplified as well, and are usually stronger than the watermark signal. All these problems make reliable scale-detection of DIVX video difficult, and often impossible. Without reliable scale detection, watermark detection is not possible.

An object of embodiments of the present invention is to at least alleviate the above mentioned problems experienced with prior art detection systems, and provide a better watermark detection system for use with highly compressed video or other multimedia data.

A further object of embodiments of the present invention is to allow the
5 performance of a more reliable scale detection process before watermark detection is carried out.

According to the present invention, there is provided a method of selecting data for use in decoding an embedded watermark in compressed multimedia data, comprising the steps of:

- 10 • calculating a quality metric for a given part of the compressed multimedia data based on the degree of compression of the multimedia data;
- including in a watermark decoding process, the given part, if its quality metric is higher than a certain threshold, and;
- excluding from the watermark decoding process, the given part, if its quality metric is
15 lower than the threshold.

Preferably the method further includes the step of using the same quality metric to select data to use in a scale-detection process performed before the watermark decoding process. In cases where no scaling has taken place, this will return a scale factor of 1. Otherwise, the scale-detection process will return a value which allows accumulation
20 buffers to be sized appropriately before a watermark is decoded.

Preferably, the quality metric is calculated on the basis of an analysis of a compressed data stream. Such a compressed data stream is provided by DIVX systems.

Suitably, in cases where access to the compressed data stream is possible, the quality metric may be determined on the basis of one of: Quantisation factors; the number of
25 Variable Length Codewords (VLCs) used to code a data frame; Motion Vectors.

The quality metric may also be calculated on the basis of a plurality of parameters.

Preferably, the quality metric may be calculated on the basis of an analysis of base-band data.

30 Preferably the quality metric is calculated on a measure of the energy of a frame.

The quality metric may also be calculated on the basis of a plurality of parameters.

Preferably, the given part of the data is a frame. Alternatively, part-frames may also be used.

Preferably, apparatus is provided to perform the method according to the invention.

5

For a better understanding of the present invention, and to understand how the same may be brought into effect, the invention will be described, by way of example only, with reference to the appended drawings in which:

10 Figure 1 shows a schematic representation of an embodiment of the present invention.

Figure 1 shows a schematic representation of the data flow in an embodiment of the invention. A data buffer 10 is arranged to receive an incoming data stream 110. The data stream 110 is, in a particular embodiment, a DIVX coded video data stream. Data buffer 10 operates to select all or part 120 of a frame of the incoming data stream, which is then analysed in quality metric calculator 20. Quality metric calculator operates on the data frame (or part thereof) 120 to establish a quality metric 130 of the input data frame 120. The quality metric is indicative of the likelihood of the particular frame including sufficient watermark energy to be used in the watermark decoding process. Methods of calculating the quality metric will be presented shortly.

20 The quality metric 130 is compared with a pre-defined level in threshold detector 30. If the quality metric indicates a high probability of the frame 120 including a suitable quantity of watermark energy, then the frame 120 is made available to the watermark detection process 40.

25 If, however, quality metric 130 falls below the pre-defined acceptable level, the threshold detector discards 50 the data in frame 120 and it will play no part in the watermark decoding process 40.

In this way, only data which has a higher probability of including sufficient watermark energy to enable a successful decode of the watermark to be performed is passed 30 to the watermark decoding process. The output of the watermark decode process is watermark 140. Alternatively, the output 140 could be a binary signal indicating either a correct decode or that no watermark was detected.

In order to determine a quality metric (Q), one or more characteristics of the data is assessed or measured. The following examples highlight attributes which may be used

in some situations. The skilled man will be aware of other attributes which may form the basis of a quality metric calculation in other situations.

The quality metric (Q) effectively provides a measure of how much the subject data has been compressed. The more compressed the data, the harder it is to extract the watermark from it.

If access to the compressed data stream is possible, there are several parameters available from the stream itself which may be used in order to determine a quality metric (Q). Some suitable parameters are:

- Quantisation Factors
- The number of Variable Length Codewords (VLCs) used to code a frame
- Motion Vectors

In a system where access to the compressed data stream is possible, a quality metric may be derived by counting the number of VLCs used to code a frame. In this case, only frames coded with more than 5000 coefficients are folded and used in the watermark detection process.

In many instances, however, access to the original compressed stream is not possible and only access to the base-band video signals is possible, for example. In such instances, access to the previously mentioned parameters is not possible and so different measures may be used to determine Q. One such measure is:

- A measure of Energy. Such a measure can be obtained, for example, by 8x8 DCT transforming blocks of a frame, quantise the coefficients with a coarse standard MPEG Quantisation matrix, and count the number of non-zero coefficients. The non-zero coefficients of a block are indicative for its energy content. If there are many high coefficients around DC frequency, this indicates that there are sharp edges in the block. A lot of non-zero coefficients means that the block has a complex structure. If there are no AC coefficients, this means that the block is flat. In general, the more non-zero coefficients there are, the more watermark energy there is likely to be available in the block.

Once a suitable quality metric (Q) has been calculated from one or more given attributes of the signal, it is possible to establish a threshold for a particular value of Q, such that data frames (or parts thereof) having a value of Q which falls below the threshold, can be discarded for the purpose of decoding an embedded watermark. The actual data frame (or part thereof) is of course retained so that its inherent data content (e.g. video) can be decoded.

The establishment of a threshold depends on the particular attribute of the data signal which was chosen as the basis of the quality metric, and may best be determined in a particular case by experimentation.

As stated previously, a further problem arises when the compressed video
5 signal has been scaled. Before the watermark can be decoded from the compressed signal, the original scale of the signal has to be recovered.

Embodiments of the present invention operate to recover scale information in a similar way to that just described to recover watermark information. To recover scale information, two accumulation buffers are correlated, with the resultant correlation giving a
10 direct indication of the scale factor.

In order to improve the results of the correlation process, the same quality metric (Q) calculated above can be used to identify candidate frames (or parts thereof) which are less heavily compressed, and thus have a higher Q . These candidate frames can be used for the scale-determining correlation process in preference to frames (or parts thereof) which
15 are more heavily compressed, and thus have a lower Q .

Experiments have shown that the scale detection process is greatly improved by being selective about which data samples are used in the correlation process. In cases where the correlation peaks would otherwise be below a defined detection threshold using prior art methods, making scale detection impossible, it is found that embodiments of the
20 invention are able to determine scale factors by selectively discarding certain data samples which do not contribute to a successful correlation.

In effect, the same technique may be used firstly to discover the scale factor of the compressed signal, which can then be used to scale the accumulation buffer appropriately and, secondly, to enable a more reliable watermark decode to take place.

Embodiments of the invention may be implemented using suitably conditioned or programmed hardware. Such hardware may include specialised hardware such as a custom ASIC, or a more general processor or DSP including operating according to a suitable
25 program.

The skilled man will be aware of other parameters which may be used as the
30 basis for calculating a quality metric, and the examples illustrated herein are not intended to limit the scope of the present invention, which is to be determined by the appended claims.